

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 03-080782

(43)Date of publication of application : 05.04.1991

(51)Int.Cl.

H04N 5/782
G06F 15/62
G10L 3/00

(21)Application number : 01-219111

(71)Applicant : NIPPON TELEGR & TELEPH CORP
<NTT>

(22)Date of filing : 24.08.1989

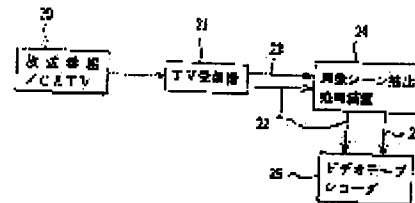
(72)Inventor : TONOMURA YOSHINOBU
KISHIMOTO TOMIO
ABE SHINJI
SUZUKI HAJIME
MASAMITSU MINEO
KIMURA YOSHIMASA

(54) SCENE EXTRACTION PROCESSING METHOD

(57)Abstract:

PURPOSE: To automatically generate the summary of a video by analyzing a voice reproduced following the reproduction of the video, comparing feature information obtained from an analytic result with the feature information with respect to the voice set in advance, and extracting a scene satisfying a condition from the video.

CONSTITUTION: An image scene extraction device 24 performs the analysis of an audio signal, and captures comparatively wide spectrum feature and the audio features of a signal level with steep rise and continuing in a short time as the audio feature with high correlation, for example, when a shout of joy arise. And it is compared with the conditions set at those features, and when it satisfies the condition, the recording of a video signal and the audio signal are performed on a video tape recorder 25. By repeating such operation, the string of highlight scenes when the shout of joy arises is recorded on the video tape recorder 25. After a program is completed, a user can observe the highlight scene in a short time by rewinding and reproducing the video tape recorder 25.



LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision
of rejection]

[Date of requesting appeal against examiner's
decision of rejection]

[Date of extinction of right]

⑫ 公開特許公報(A)

平3-80782

⑤Int.Cl.⁵

H 04 N 5/782
G 06 F 15/62
G 10 L 3/00

識別記号

3 4 0 A
3 0 1 F

庁内整理番号

7734-5C
8125-5B
8842-5D

④3公開 平成3年(1991)4月5日

審査請求 未請求 請求項の数 2 (全4頁)

⑭発明の名称 シーン抽出処理方法

⑰特 願 平1-219111

⑱出 願 平1(1989)8月24日

⑲発 明 者 外 村 佳 伸 東京都千代田区内幸町1丁目1番6号 日本電信電話株式会社内
⑲発 明 者 岸 本 登 美 夫 東京都千代田区内幸町1丁目1番6号 日本電信電話株式会社内
⑲発 明 者 安 部 伸 治 東京都千代田区内幸町1丁目1番6号 日本電信電話株式会社内
⑲発 明 者 鈴 木 元 東京都千代田区内幸町1丁目1番6号 日本電信電話株式会社内
⑲出 願 人 日本電信電話株式会社 東京都千代田区内幸町1丁目1番6号
⑲代 理 人 弁理士 秋田 収喜
最終頁に続く

明 細 書

1. 発明の名称

シーン抽出処理方法

2. 特許請求の範囲

1. 音声情報を伴う一連の映像の中から部分的なシーンを抽出するシーン抽出処理方法において、映像再生時に付随して再生される音声解析し、該解析結果から得られる特徴情報と予め設定された音声に関する特徴情報とを比較し、条件を満足するシーンを該映像から抽出することを特徴とするシーン抽出処理方法。

2. 請求項1記載のシーン抽出処理方法で得られる抽出シーンをつなぎ合わせて新たな映像列を構成することを特徴とするシーン抽出処理方法。

3. 発明の詳細な説明

〔産業上の利用分野〕

本発明は、音声の付随するビデオ画像等の映像から重要なシーンを抽出することにより、映像の要約を作成するシーン抽出処理方法に関する。

〔従来技術〕

従来、映像からのシーン抽出に関する技術としては、映像情報そのものを解析し、映像情報の変化をとらえることにより自動的にカット変化点を検出する技術がある(例えば、清水他、「映像カット変化点検出器」、1988年テレビジョン学会全国大会、9-4、1988)。

第3図は、前記した従来方式を説明するためのブロック図である。端子30に入力されたビデオ信号はフレーム単位でアナログ・デジタル変換(A/D変換31)されて、フレームメモリ32に記憶され、さらに次の入力フレームの画信号と画素毎に画素比較部33で比較され、結果としてフレーム間での差分としての画素数が得られる。得られた差分は変化判断部34において、あらかじめ設定されている判断しきい値35と比較され、しきい値を越えた場合にはカット変化有りとして、しきい値を越えない場合にはカット変化なしとして、端子36から結果が出力される。

またこの他、音声のみに対する解析処理と応用に関しては、テープレコーダーなどで従来より一

般的に行われている方法として、音声レベルの極めて低い部分が無音部分と判断することにより、音声をセグメンテーション化する方法がある。

〔発明が解決しようとする課題〕

しかしながら、上記した前者の従来技術は、映像情報処理のみによってカット変化点を検出しているため、映像情報に変化さえあれば変化点検出となり、映像情報の意味的な表現内容が変化しなくてもカット変化ありと判断するため、映像情報の内容に関わる特定のシーンを抽出することは難しいという問題があった。

また、編集された映像が処理対象の場合、一般に数秒毎に映像の変化点が見られるため、検出変化点が非常に多くなってしまいうとともに、検出結果を例えばそのまま映像の要約作成に用いようとしても、ほとんど元の映像に近い映像になってしまうという問題があった。

更に、上記した後者の従来技術は、音声の有無のみを対象としているため、音声のセグメンテーション程度以上のことはできないという問題があ

った。

本発明の目的は、音声の付随する映像情報から重要なシーン抽出を行い、映像要約を作成するシーン抽出方法を提供することにある。

〔課題を解決するための手段〕

前記目的を達成するために、第1の発明では、映像再生時に付随して再生される音声を解析し、該解析結果から得られる特徴情報と予め設定された音声に関する特徴情報とを比較し、条件を満足するシーンを該映像から抽出している。

そして、第2の発明では、第1の発明で得られた抽出シーンをつなぎ合わせて新たな映像列を構成している。

〔作用〕

本発明では、映像再生時に付随して再生される音声をモニタしながら解析し、音声に含まれる特徴情報を抽出する。そして、抽出した特徴量をあらかじめ設定されている条件と比べ、条件を満足するシーンを抽出する。これにより、目的とする映像の内容に関わりの深いシーンを取り出すこと

ができ、従って一連の長い動画像から重要なシーンのみによる短い要約としての映像列を構成することができる。

〔実施例〕

以下、本発明の一実施例を図面を用いて具体的に説明する。

なお、実施例を説明するための全図において、同一機能を有するものは同一符号を付け、その繰り返しの説明は省略する。

第1図は、本発明の一実施例に係る処理手順を示したフローチャートである。

ステップ1において音声入力ユニットを制御することにより音声入力を行い、ステップ2において入力音声の解析を行い、特徴抽出を行う。抽出する特徴は、例えば、音声の信号レベル、ピーク周波数、主要音声スペクトル幅等であり、かつ動画像内容に相関の高い特徴を抽出する。

ステップ3において抽出した特徴量をあらかじめ設定されている条件と比べ、条件に合う場合のみ次の処理を行う。

ステップ4においては、指示された処理が実時間記録でシーン抽出を行い要約を作る実時間記録処理か、マーキング情報として画像のフレーム番号などを記録しておくマーキング処理かを判断する。

実時間記録処理の場合はステップ5において映像の記録を開始する。マーキング処理の場合は、ステップ6においてその時点の画像フレーム番号を調べ、シーン開始フレーム番号として記録する。

次に、ステップ7において抽出処理終了条件を調べるが、条件設定は適用する応用例に合わせて設定される。例えば、一定時間経つことを終了条件としたり、あるいはステップ1～3と同様の解析／特徴抽出処理を行い、条件が不成立になることを終了条件とする場合などである。

終了条件成立時には、実時間記録処理であれば記録を停止し、またマーキング処理の場合は、終了フレーム番号としてその時点のフレーム番号を記録する。以上の処理を繰り返すことにより、対象とする映像についてシーン抽出処理を行う。

第2図は、以上で説明した処理手順を応用したスポーツ番組自動要約作成システムを示している。

20は例えばスポーツ番組として試合を中継している放送波あるいはケーブルTVからの信号、21は受像機である。24は上述したシーン抽出処理を行う画像シーン抽出処理装置であり、受像機21から出力されるビデオ信号22および音声信号23を受像機21から受け、ビデオテープレコーダー25を制御して抽出したビデオ信号、音声信号を記録する。

画像シーン抽出処理装置24は、音声信号の解析を行って、観客の歓声が上がった時に相関の高い音声特徴として、比較的幅広いスペクトル特徴、急な立ち上がりと短時間持続する信号レベルなどの音声特徴をとらえ、予めそれらの特徴に設定してある条件と比べ、条件に合う場合には実時間記録処理としてビデオテープレコーダー25にビデオ信号および音声信号の記録を開始する。

記録中も音声信号の解析/条件判断を続け、条件不成立となってもその時点からタイマーを起動し一定時間記録することとするが、その間にさら

に条件が成立する状態が発生すればその時点でタイマーをリセットし、起動する。

以上の動作を繰り返すことにより、観客の歓声の上がったハイライトシーン列がビデオテープレコーダー25に記録される。番組終了後、利用者が該ビデオテープレコーダー25を巻戻し、再生すれば短時間にハイライトシーンを見ることができる。

以上の処理は、実時間記録処理に応用した例であるが、予め番組分のビデオ信号、音声信号を全部記録しておき、画像シーン抽出処理装置24の処理の中でマーキング情報のみ格納しておき、利用者再生時に開始フレームと終了フレームのマーキングのある区間の再生を行うことも可能である。

以上、本発明を実施例にもとづき具体的に説明したが、本発明は、前記実施例に限定されるものではなく、その要旨を逸脱しない範囲において種々変更可能であることは言うまでもない。

〔発明の効果〕

以上、説明したように、本発明によれば、音声信号の解析/特徴抽出を行い、映像信号と相関の

高い部分を検出することにより、映像中の内容に関わるシーン抽出を自動的に行うことができるとともに、それらをつなぎ合わせるにより映像の要約を自動作成することができる。

また、放送番組等の自動要約作成に利用することにより、長時間の映像を利用者は時間を短縮して見ることができる。

さらに、映像ライブラリなどでは映像情報のブラウズ（流し見）用の映像を自動的に作成できるので、省力化されるという効果もある。

4. 図面の簡単な説明

第1図は、本発明の一実施例に係る処理手順を示したフローチャート、

第2図は、本発明の処理手順を応用したスポーツ番組自動要約作成システム、

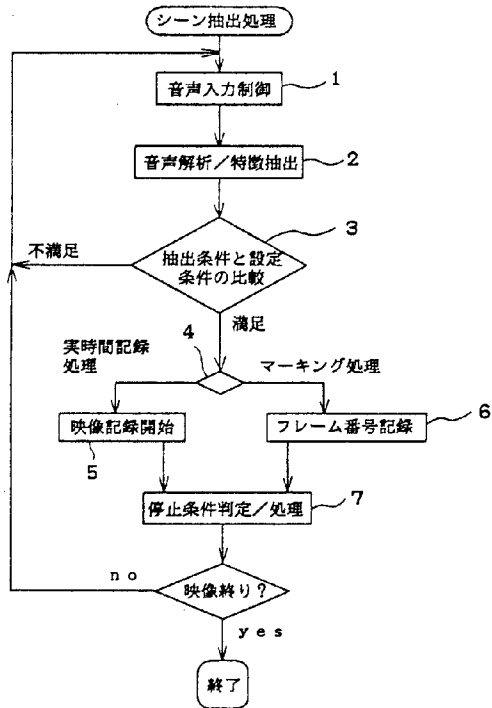
第3図は、従来方法によるカット検出器のブロック図を示す。

図中、21…TV受像機、22…ビデオ信号、23…音声信号、24…画像シーン抽出処理装置、25…ビデオテープレコーダー、31…A/D変換、32…フ

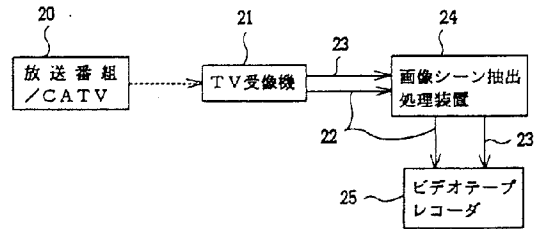
レームメモリ、33…画像比較部、34…変化判断部、35…判断しきい値。

代理人 弁理士 秋田収喜

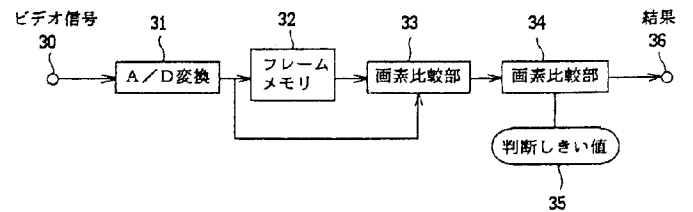
第1図



第2図



第3図



第1頁の続き

⑦2発 明 者 正 満 峰 夫

東京都千代田区内幸町1丁目1番6号 日本電信電話株式会社内

⑦2発 明 者 木 村 義 政

東京都千代田区内幸町1丁目1番6号 日本電信電話株式会社内